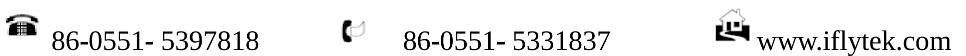


XF-S4040 中文语音合成芯片
数据手册

安徽科大讯飞信息科技股份有限公司

安徽省合肥市国家级高新技术产业开发区信息产业基地讯飞语音大厦


86-0551- 5397818　86-0551- 5331837　www.iflytek.com

版本历史

版 本	日 期	修改记录	作 者
Rev0.1	2008-3-7	非正式发布初稿	科大讯飞
Rev0.2	2008-7-28	修改相关外观图和引脚说明	科大讯飞
Rev0.3	2008-7-29	1.修改公司名称 2.修改音频参考电路	科大讯飞

声　明

科大讯飞保证本手册提供信息的准确性和可靠性，但并不对文本中可能出现的文字错误或疏漏负责。科大讯飞保留更改本手册的权利，如有修改，恕不相告。请在订购时联系我们以获得产品最新信息。对任何用户使用我们产品时侵犯第三方版权或其他权利的行为科大讯飞概不负责。另外，在科大讯飞未明确表示产品有该项用途时，对于产品使用在极端条件下导致一些可预见失灵或损毁而造成的损失概不负责。

目　　录

1. 概述 - 4 -
2. 特点 - 4 -
3. 主要应用领域 - 5 -
4. 系统框图 - 5 -
5. 功能描述 - 6 -
6. 引脚说明 - 7 -
6.1 XF-S4040-CN 中文语音合成芯片 - 7 -
6.1.1 芯片封装 - 7 -
6.1.2 引脚位置 - 7 -
6.1.3 引脚信号说明 - 8 -
6.2 XF-S4040M 中文语音合成模块 - 11 -
6.2.1 XF-S4040M 引脚图 - 11 -
6.2.2 XF-S4040M 引脚定义 - 12 -
7. 通讯方式 - 12 -
7.1 串口通讯模式 - 13 -
7.1.1 通讯框图 - 13 -
7.1.2 通讯传输字节格式 - 13 -
7.2 I2C 通讯模式 - 14 -
7.2.1 通讯框图 - 14 -
8. 芯片控制 - 15 -
8.1 控制命令 - 15 -
8.2 芯片反馈 - 15 -
8.3 芯片状态查询 - 15 -
9. 电路参考 - 16 -
9.1 复位方式 - 16 -
9.2 电源接法 - 16 -
9.2.1 单 5v 供电 - 16 -
9.2.2 3.3v 和 1.8v 两路供电 - 16 -
9.3 与上位机通讯连接 - 17 -
9.3.1 UART 通讯 - 17 -
9.3.2 I2C 通讯 - 18 -
9.4 音频功放电路 - 18 -
10. 极限值 - 19 -
11. XF-S4040 语音合成芯片通讯协议 - 20 -
11.1 上位机发送的控制命令 - 20 -

11.2 芯片反馈信息 - 25 -
11.3 芯片工作状态引脚输出 - 26 -
11.4 上位机对 XF-S4040 芯片调用方式 - 26 -
11.4.1 简单调用 - 26 -
11.4.2 标准调用方式 - 26 -

图表目录

图表 1 系统框图 - 5 -
图表 2 XF-S4040 封装引脚图 - 7 -
图表 3 XF-S4040-CN 中文语音合成芯片引脚位置图 - 7 -
图表 4 XF-S4040M 引脚图 - 11 -
图表 5 XF-S4040M 引脚说明 - 12 -
图表 6 XF－S4040 串口通讯连接示意图 - 13 -
图表 7 串口通讯传输字节格式 - 13 -
图表 8 XF-S4040 I2C 总线挂接 - 14 -
图表 9 I^2C 总线时序图 - 14 -
图表 10 控制命令简表 - 15 -
图表 11 芯片反馈信息表 - 15 -
图表 12 XF-S4040 采用 5V 供电的参考电路 - 16 -
图表 13 XF-S4040 采用 3.3V 和 1.8V 供电的参考电路 - 16 -
图表 14 XF-S4040 芯片与 PC 的异步串口连接示意图 - 17 -
图表 15 XF-S4040 芯片与 MCU 的异步串口连接示意图 - 17 -
图表 16 XF-S4040 芯片与 MCU 的 I2C 接口连接示意图 - 18 -
图表 17 XF-S4040 芯片与音频功放 LM4665 的接口 - 18 -
图表 18 XF-S4040 芯片与 LM386 的接口示意图 - 19 -
图表 19 电气极限参数 - 19 -
图表 20 静态特性 - 19 -
图表 21 命令帧封装格式 - 20 -
图表 22 命令帧说明 - 20 -
图表 23 数据区命令字和命令参数 - 20 -
图表 24 语音合成命令 - 22 -
图表 25 停止合成命令 - 23 -
图表 26 暂停合成命令 - 23 -
图表 27 恢复合成命令 - 24 -
图表 28 系统状态查询命令 - 24 -
图表 29 进入休眠状态命令 - 25 -
图表 30 芯片反馈信息 - 25 -
图表 31 工作状态引脚描述 - 26 -

1. 概述

语音合成技术，简称 TTS（Text To Speech）技术，该技术解决如何将文字信息转化为声音信息的问题，从而变看为听，使得人们获取信息的方式更加丰富和自然。语音合成技术是实现人机语音通信的关键技术之一，涉及声学、语言学、数字信号处理、计算机科学等多种学科知识，是中文信息处理领域的一项前沿技术。

XF-S4040 中文语音合成芯片是面向中高端应用的语音合成芯片，可通过异步串口、I^2C 接口接收任意文本，并将文本转化为语音输出。

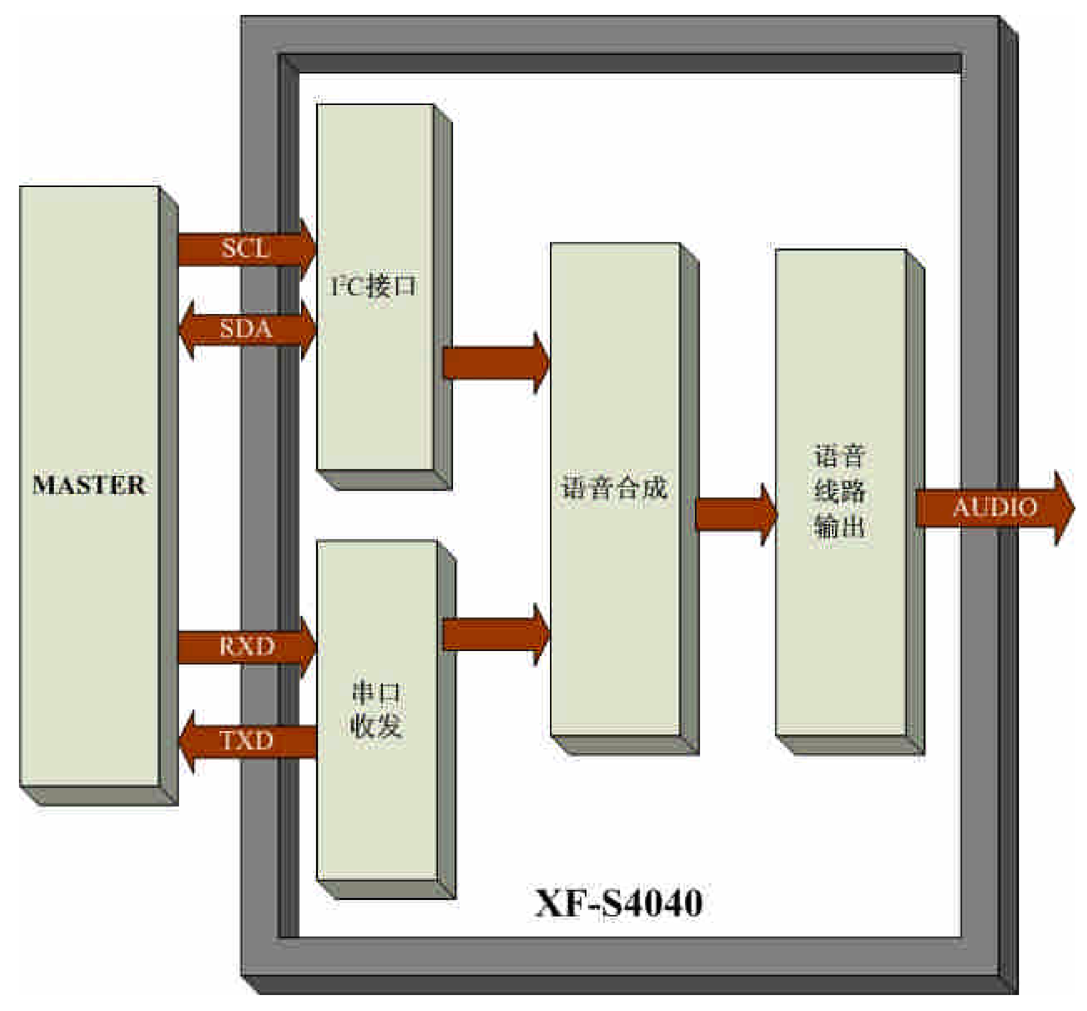
2. 特点

- 可合成任意的中文文本，支持英文字母的合成；
- 支持 GB18030、BIG5、UNICODE、UTF-8 四种内码格式的文本；
- 具有智能的文本分析处理算法，可正确的识别和处理数值、号码、时间日期及一些常用的度量衡符号，具备较强多音字处理和中文姓氏处理能力；
- 提供六种发音人：其中女性发音人两种，男性发音人两种，童声一种，卡通效果音色一种；
- 清晰、自然、准确的文语音合成效果；
- 集成提示音效；
- 支持多种控制命令，包括：合成、停止、暂停合成、继续合成等；
- 支持多种文本控制标记，提升文本处理的正确率；
- 支持休眠功能，在休眠状态下可降低功耗；
- 支持 UART、I^2C 两种数据通讯接口；
- Line out 音频输出；

3. 主要应用领域

- 车载 GPS 语音导航系统
- 车载电话号码、短信播报系统
- 公交车语音报站、公告播报系统
- 智能仪表
- 自动售货
- 语音导游
- 排队机

4. 系统框图



图表 1 系统框图

5. 功能描述

XF-S4040 语音合成芯片在文本处理前端集成了多种智能的文本分析算法，将规则和统计等处理方法进行了有机的结合，具备了较强的分词、多文本内码识别、多音字识别、韵律环境预测能力，可以自动识别和处理日期时间、数值金额、电话号码等文本的能力，可以利用特殊的控制标识对芯片的合成效果进行语速、语调和音量上的调整，以及利用控制标识对姓氏进行准确的识别。

- 文本合成功能：芯片支持任意中文文本的合成，可以采用 GB18030、BIG5、Unicode 和 UTF8 编码方式，其中 GB18030 是 GB2312 和 GBK 的取代版本； BIG5 是台湾地区常用的文本编码格式； Unicode 编码是国际通用的编码格式；UTF-8 编码是一种被广泛应用的编码，这种编码致力于把全球的语言纳入一个统一的编码，目前已经将几种亚洲语言纳入。芯片在默认情况下按 GB18030 编码格式处理中文文本。芯片支持英文字母的合成，遇到英文单词时按字母方式发音。

- 文本分析算法：芯片具有智能的文本分析算法，对常见的数值、电话号码、时间日期、度量衡符号等格式的文本，系统能够根据内置的文本匹配规则进行正确的识别和处理。如“2004/12/21”可以识别并读作 “二零零四年十二月二十一日”。

- 提示音：芯片内带有音效提示音，可用于铃声和信息提醒。

- 文本标记功能：芯片允许在文本中插入特殊控制标记，利用这些控制标记可以调节芯片的参数，提升芯片文本分析的正确率，如：标记句子的韵律、产生静音间隔、标识中文姓氏文本、支持合成拼音等。

- 简易规范的开发接口：芯片采用的通信方式，命令协议简练有效；可通过统一的“合成命令”接口播放音效和中文文本，甚至进行系统参数的设置。开发者只需要掌握少量命令便可轻松上手，使用所有主要功能。

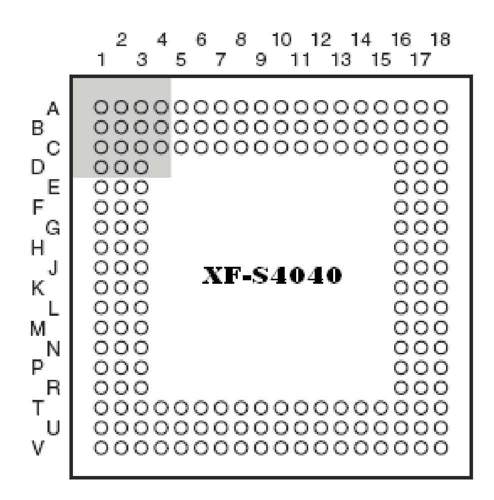
- 丰富的通讯模式：芯片支持 UART、I^2C 两种通讯方式。解决了用户串口资源不足的问题。

- 多种工作模式：芯片支持两种工作模式，分别是：正常工作模式、休眠模式。芯片提供休眠命令帧，芯片接收到此命令后立即休眠状态，此时必需对芯片复位方能使芯片进入正常工作模式。

6. 引脚说明

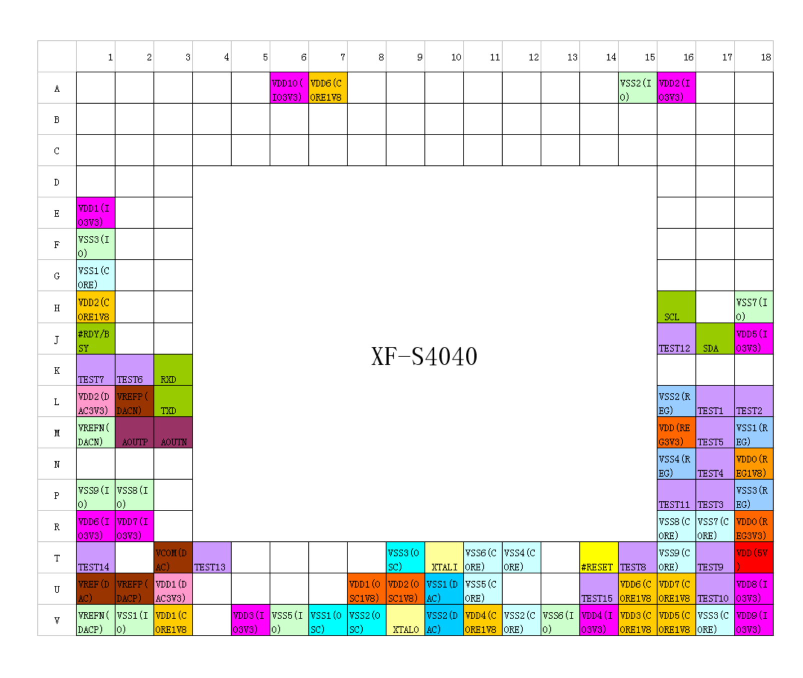
6.1 XF-S4040-CN 中文语音合成芯片

6.1.1 芯片封装



图表 2 XF-S4040 封装引脚图

6.1.2 引脚位置



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
A						VDD10(IO3V3)	VDD6(CORE1V8								VSS2(IO)	VDD2(IO3V3)		
B																		
C																		
D																		
E	VDD1(IO3V3)																	
F	VSS3(IO)																	
G	VSS1(CORE)																	
H	VDD2(CORE1V8															SCL		VSS7(IO)
J	#RDY/BSY								XF-S4040							TEST12	SDA	VDD5(IO3V3)
K	TEST7	TEST6	RXD															
L	VDD2(DAC3V3)	VREFP(DACN)	TXD													VSS2(REG)	TEST1	TEST2
M	VREFN(DACN)	AOUTP	AOUTN													VDD(REG3V3)	TEST5	VSS1(REG)
N																VSS4(REG)	TEST4	VDDO(REG1V8)
P	VSS9(IO)	VSS8(IO)														TEST11	TEST3	VSS3(REG)
R	VDD6(IO3V3)	VDD7(IO3V3)														VSS8(CORE)	VSS7(CORE)	VDDO(REG3V3)
T	TEST14		VCOM(DAC)	TEST13					VSS3(OSC)	XTALI	VSS6(CORE)	VSS4(CORE)		#RESET	TEST8	VSS9(CORE)	TEST9	VDD(5V)
U	VREF(DAC)	VREFP(DACP)	VDD1(DAC3V3)					VDD1(OSC1V8)	VDD2(OSC1V8)	VSS1(DAC)	VSS5(CORE)			TEST15	VDD6(CORE1V8	VDD7(CORE1V8	TEST10	VDD8(IO3V3)
V	VREFN(DACP)	VSS1(IO)	VDD1(CORE1V8		VDD3(IO3V3)	VSS5(IO)	VSS1(OSC)	VSS2(OSC)	XTALO	VSS2(DAC)	VDD4(CORE1V8	VSS2(CORE)	VSS6(IO)	VDD4(IO3V3)	VDD3(CORE1V8	VDD5(CORE1V8	VSS3(CORE)	VDD9(IO3V3)

图表 3 XF-S4040-CN 中文语音合成芯片引脚位置图

6.1.3 引脚信号说明

表格 1 文本特殊标记说明

	引脚信号	引脚	类型	信号说明
振荡器	#RESET	T14	I	复位，低电平有效
	XTALI	T10	I	振荡输入
	XTALO	V9	O	振荡输出
	VDD1(OSC1V8)	U8	P	振荡电源 1.8 V
	VDD2(OSC1V8)	U9	P	振荡电源 1.8 V
	VSS1(OSC)	V7	I	振荡地
	VSS2(OSC)	V8	P	振荡地
	VSS3(OSC)	T9	P	振荡地
接　口	RXD	K3	I	串口输入
	TXD	L3	O	串口输出
	SCL	H16	I	I^2C 时钟输入，漏极开路
	SDA	J17	I	I^2C 数据输入/输出，漏极开路
	#RDY/BSY	J1	O	工作状态指示 高电平：忙，低电平：就绪
数模转换	AOUTP	M2	O	差分音频输出正端
	AOUTN	M3	O	差分音频输出负端
	VCOM(DAC)	T3	RV	数模转换公共端基准电源
	VREF(DAC)	U1	RV	数模转换基准电源
	VREFP(DACP)	U2	RV	差分音频正端基准电源
	VREFP(DACN)	L2	RV	差分音频负端基准电源
	VREFN(DACP)	V1	RV	差分音频正端基准地
	VREFN(DACN)	M1	RV	差分音频负端基准地
	VDD1(DAC3V3)	U3	P	数模转换电源
	VDD2(DAC3V3)	L1	P	数模转换电源
	VSS1(DAC)	V10	P	数模转换地
	VSS2(DAC)	U10	P	数模转换地
电压调节器	VDD(5V)	T18	P	5V 电源输入
	VDD(REG3V3)	M16	P	电压调节器电源
	VDDO(REG3V3)	R18	P	3.3V 电源调节器输出
	VDDO(REG1V8)	N18	P	1.8V 电源调节器输出
	VSS1(REG)	M18	P	电压调节器地
	VSS2(REG)	L16	P	电压调节器地
	VSS3(REG)	P18	P	电压调节器地
	VSS4(REG)	N16	P	电压调节器地
内核电源	VDD1(CORE1V8)	V3	P	内核电源 1.8 V

	VDD2(CORE1V8)	H1	P	内核电源 1.8 V
	VDD3(CORE1V8)	V15	P	内核电源 1.8 V
	VDD4(CORE1V8)	V11	P	内核电源 1.8 V
	VDD5(CORE1V8)	V16	P	内核电源 1.8 V
	VDD6(CORE1V8)	U15	P	内核电源 1.8 V
	VDD7(CORE1V8)	U16	P	内核电源 1.8 V
	VSS1(CORE)	G1	P	内核地
	VSS2(CORE)	V12	P	内核地
	VSS3(CORE)	V17	P	内核地
	VSS4(CORE)	T12	P	内核地
	VSS5(CORE)	U11	P	内核地
	VSS6(CORE)	T11	P	内核地
	VSS7(CORE)	R17	P	内核地
	VSS8(CORE)	R16	P	内核地
	VSS9(CORE)	T16	P	内核地
接口电源	VDD1(IO3V3)	E1	P	接口电源 3.3 V
	VDD2(IO3V3)	A16	P	接口电源 3.3 V
	VDD3(IO3V3)	V5	P	接口电源 3.3 V
	VDD4(IO3V3)	V14	P	接口电源 3.3 V
	VDD5(IO3V3)	J18	P	接口电源 3.3 V
	VDD6(IO3V3)	R1	P	接口电源 3.3 V
	VDD7(IO3V3)	R2	P	接口电源 3.3 V
	VDD8(IO3V3)	U18	P	接口电源 3.3 V
	VDD9(IO3V3)	V18	P	接口电源 3.3 V
	VDD10(IO3V3)	A7	P	接口电源 3.3 V
	VSS1(IO)	V2	P	接口地
	VSS2(IO)	A15	P	接口地
	VSS3(IO)	F1	P	接口地
	VSS4(IO)	A6	P	接口地
	VSS5(IO)	V6	P	接口地
	VSS6(IO)	V13	P	接口地
	VSS7(IO)	H18	P	接口地
	VSS8(IO)	P2	P	接口地
	VSS9(IO)	P1	P	接口地
测　试	TEST1	L17	I	测试引脚，悬空
	TEST2	L18	I	测试引脚，悬空
	TEST3	P17	O	测试引脚，悬空
	TEST4	N17	O	测试引脚，悬空
	TEST5	M17	O	测试引脚，悬空

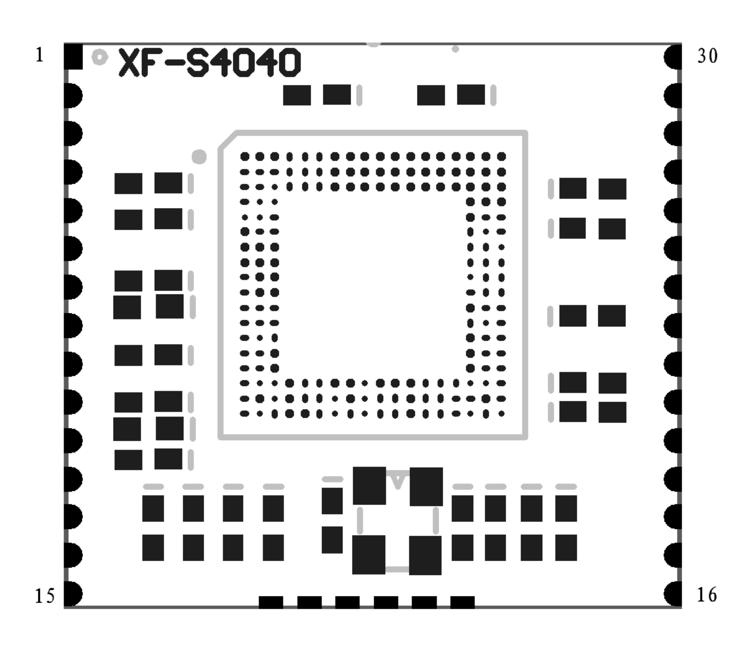
	TEST6	K2	O	测试引脚，悬空
	TEST7	K1	O	测试引脚，悬空
	TEST8	T15	O	测试引脚，悬空
	TEST9	T17	I	测试引脚，悬空
	TEST10	U17	I	测试引脚，悬空
	TEST11	P16	O	测试引脚，悬空
	TEST12	J16	I	测试引脚，悬空
	TEST13	T4	I	测试引脚，悬空
	TEST14	T1	I	测试引脚，悬空
	TEST15	U14	I	测试引脚，悬空

注：

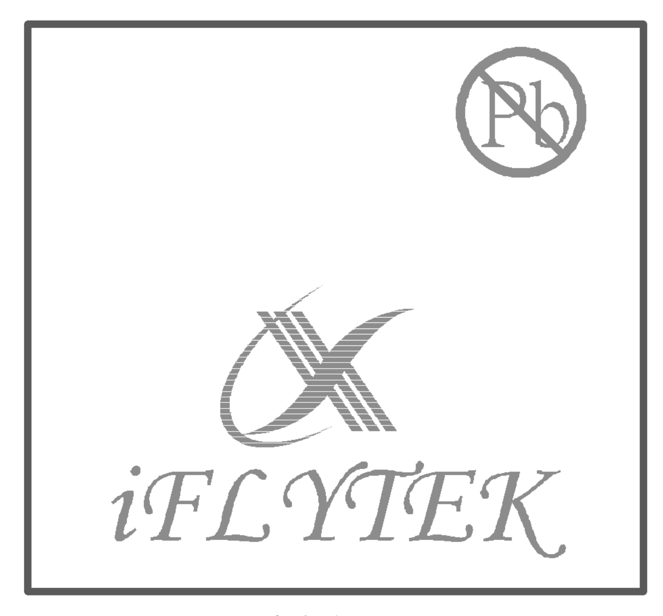
I：输入引脚
O：输出引脚
RV：基准电源
P：电源引脚

6.2 XF-S4040M 中文语音合成模块

6.2.1 XF-S4040M 引脚图



顶部视图



底部视图

图表 4 XF-S4040M 引脚图

6.2.2 XF-S4040M 引脚定义

图表 5 XF-S4040M 引脚说明

引脚	信号	描述	引脚	信号	描述
1	SCL	I^2C 时钟输入，漏极开路	30	TEST12	悬空
2	SDA	I^2C 数据输入/输出，漏极开路	29	TEST1	悬空
3	#RESET	复位，低电平有效	28	TEST2	悬空
4	TXD	串口输出	27	VDD(CORE1V8)	内核电源 1.8 V
5	RXD	串口输入	26	VDDO(REG1V8)	1.8V 电源调节器输出
6	#RDY/BSY	工作状态指示（高电平：忙，低电平：就绪）	25	TEST4	悬空
7	TEST6	悬空	24	TEST5	悬空
8	TEST7	悬空	23	TEST3	悬空
9	AOUTP	差分音频输出正端	22	VDDO(REG3V3)	3.3V 电源调节器输出
10	AOUTN	差分音频输出负端	21	VDD(IO3V3)	接口电源 3.3 V
11	TEST13	悬空	20	VDD(5V)	5V 电源输入
12	TEST14	悬空	19	TEST9	悬空
13	VDD(DACREF)	模拟 3.3V 参考电压输入	18	TEST10	悬空
14	VDD(DAC3V3)	模拟 3.3V 输入	17	TEST15	悬空
15	GND	电源地	16	GND	电源地

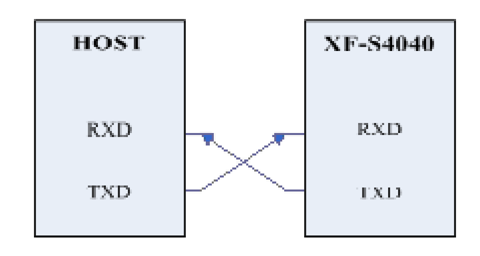
7. 通讯方式

XF-S4040 支持异步串口(UART)和 I^2C 通讯方式，通过串口和 I^2C 接受上位机发送的命令

和数据，允许发送数据的最大长度为 4K 字节。

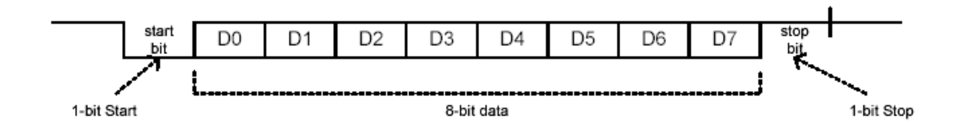
7.1 串口通讯模式

7.1.1 通讯框图



图表 6 XF－S4040 串口通讯连接示意图

7.1.2 通讯传输字节格式



图表 7 串口通讯传输字节格式

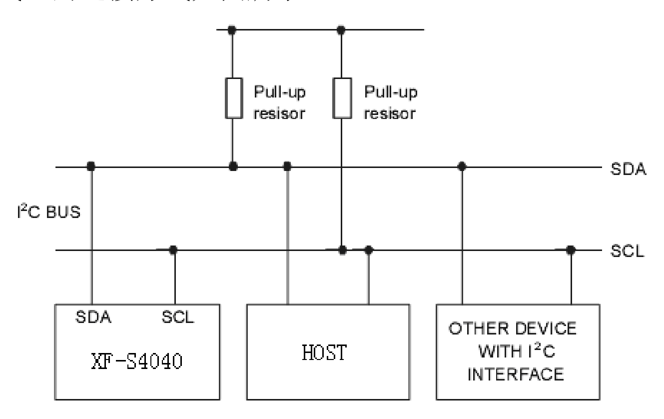
1） 通讯标准：UART
2） 波特率：9600 bps（默认值）
3） 起始位：1bit
4） 数据位：8 bits
5） 停止位：1 bit
6） 校验：无

7.2 I²C 通讯模式

7.2.1 通讯框图

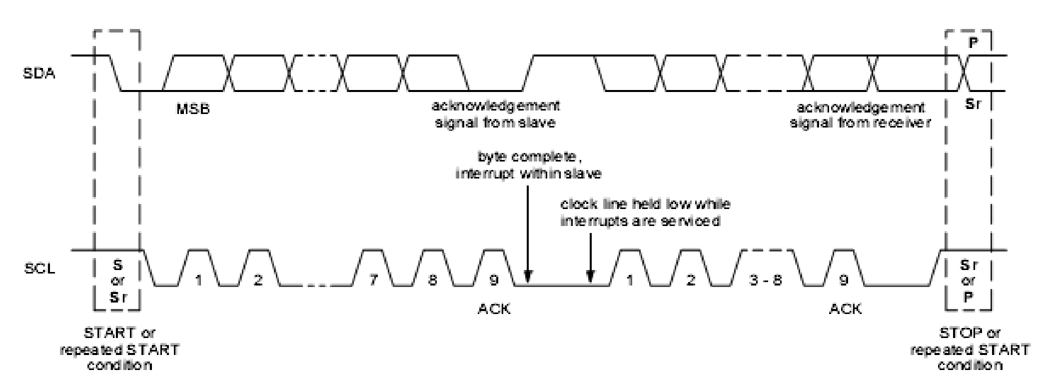
I²C 总线是一个为简化多个智能设备之间相互传输信息而设计的两线、双向网络。挂接在 I²C 总线上的设备只使用两条线（SCL 和 SDA）来实现双向通讯。I²C 总线上能挂接多个设备，由于使用 7bits 数据作为设备地址，所以总线上最多能挂接 127 个设备。

XF-S4040 采用标准 I²C 总线接口，设置为 Slave 端（从设备）。**XF-S4040 芯片的默认设备地址为 0x40**。XF-S4040 的 I²C 地址采用 7bit 描述，在发送完起始位后，紧接着发送 7bit 地址和 1bit 命令（0 代表主设备向从设备写数据，1 代表主设备向从设备读数据），因此当主设备向从设备写数据时候，发送参数为：0x80；当主设备向从设备读数据的时候，发送参数为：0x81。在 I²C 总线上的连接方式如图所示：



图表 8 XF-S4040 I2C 总线挂接

I²C 总线通讯的时序如图所示，传输的数据位数是 8bits，每次传输数据的字节数无限制，每一个字节传输后返回确认位。



图表 9 I²C 总线时序图

XF-S4040 采用的标准 I²C 总线接口，关于 I²C 总线通讯的详细信息，**请参考"The I²C-bus specification"**。

8. 芯片控制

8.1 控制命令

XF-S4040 提供了多种控制命令，芯片的控制命令简列如下：

命令功能	说明
合成命令	合成本次发送的文本
停止合成命令	停止当前的合成动作
暂停合成命令	暂停正在进行的合成
恢复合成命令	继续合成被暂停的文本
状态查询命令	查询当前芯片的工作状态
休眠命令	芯片进入休眠节能状态，如要唤醒需要重新复位芯片

图表 10 控制命令简表

注：当芯片正在合成文本的时候，如果又接收到一帧有效的合成命令，芯片会立即停止当前正在合成的文本，转而合成最新接收到的文本。

8.2 芯片反馈

针对控制命令帧，XF-S4040 芯片会向上位机发送单字节的反馈信息，上位机可根据这个反馈来判断芯片目前的工作状态。

图表 11 芯片反馈信息表

名称	触发条件
初始化成功回传	系统初始化成功
收到正确的命令帧回传	收到正确的命令帧回传
收到不能识别命令帧回传	收到错误的命令帧回传
系统状态回传	收到状态查询命令帧回传

8.3 芯片状态查询

可以通过硬件和软件两种方式查询 XF-S4040 的工作状态。

硬件方式是通过查询输出引脚 $\overline{\text{RDY}}$/BSY 的电平，来判断芯片工作的状态。当 $\overline{\text{RDY}}$/BSY 处于地电平时，表明芯片处于就绪状态；当 $\overline{\text{RDY}}$/BSY 处于高电平状态时，表明芯片处于合成文本状态。

软件方式是通过系统状态查询帧来判断芯片工作的状态。当上位机发送系统状态查询命令帧给芯片后，芯片会立即向上位机发送当前系统状态回传。上位机根据系统状态的回传数据来判断当前芯片是处于空闲状态还是忙碌状态。

9. 电路参考

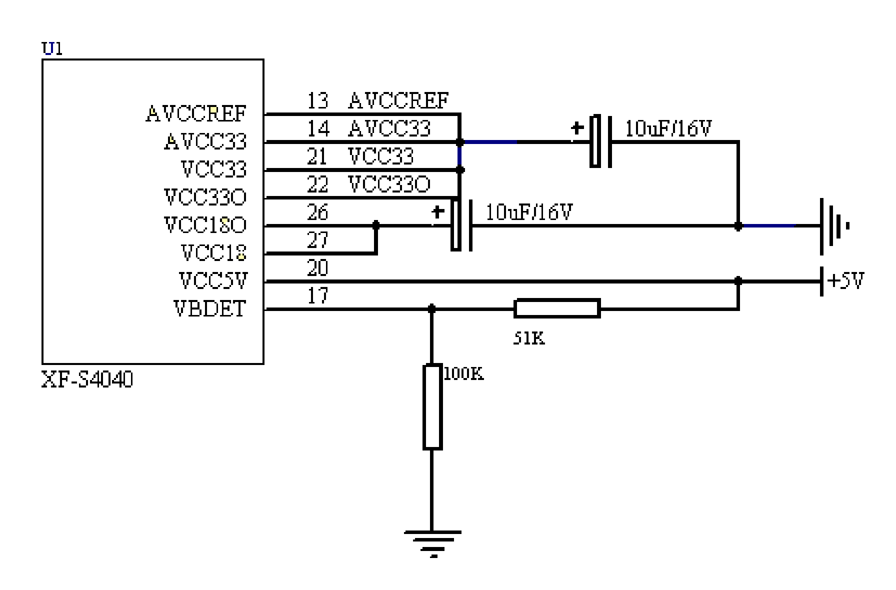
9.1 复位方式

XF-S4040 上电会自动复位，无需用户复位。如果用户需要复位，将芯片的 $\overline{RST}$ 引脚置为低电平，时间 10 微秒以上即可，然后将此引脚置为高电平，芯片将重新启动。

提醒：此引脚被置为低电平时芯片将处于复位状态。

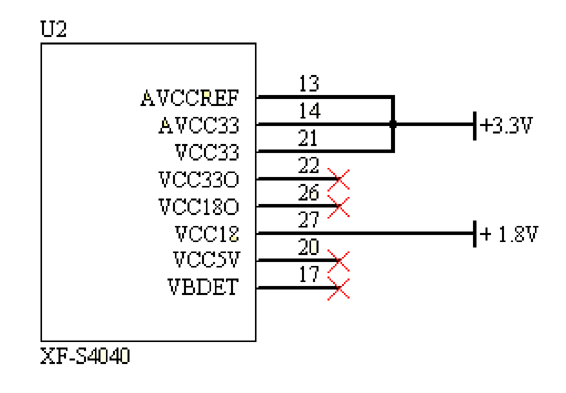
9.2 电源接法

9.2.1 单 5v 供电



图表 12 XF-S4040 采用 5V 供电的参考电路

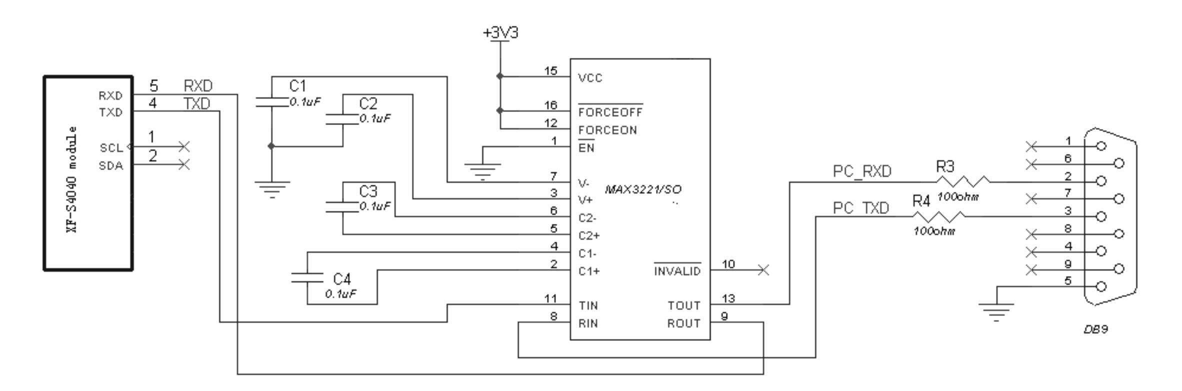
9.2.2 3.3v 和 1.8v 两路供电



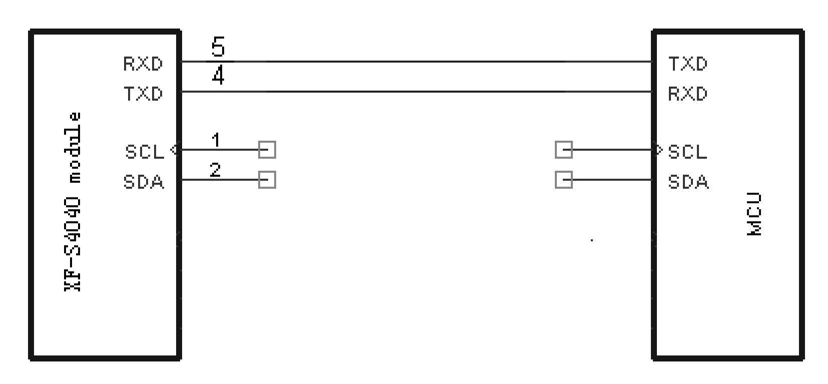
图表 13 XF-S4040 采用 3.3V 和 1.8V 供电的参考电路

9.3 与上位机通讯连接

9.3.1 UART 通讯

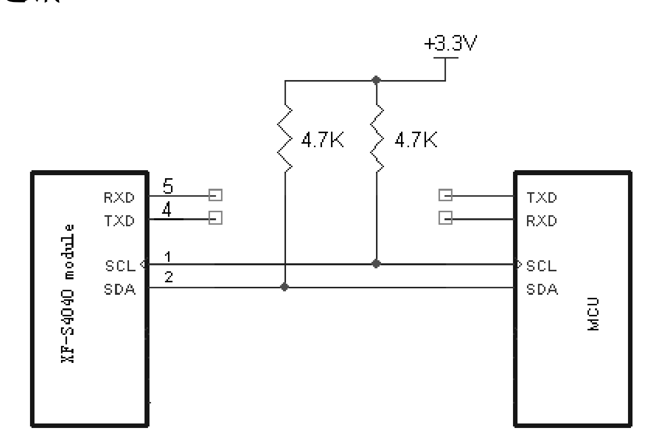


图表 14 XF-S4040 芯片与 PC 的异步串口连接示意图



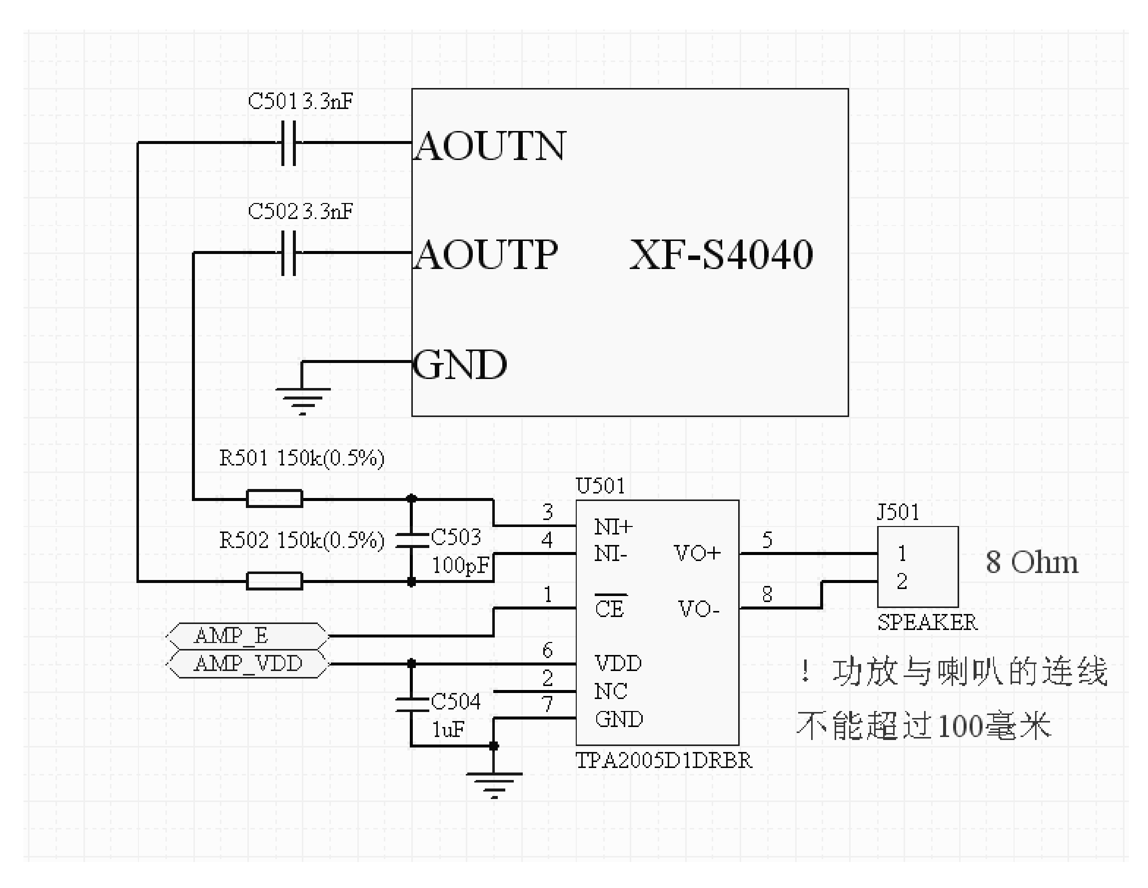
图表 15 XF-S4040 芯片与 MCU 的异步串口连接示意图

9.3.2 I²C 通讯

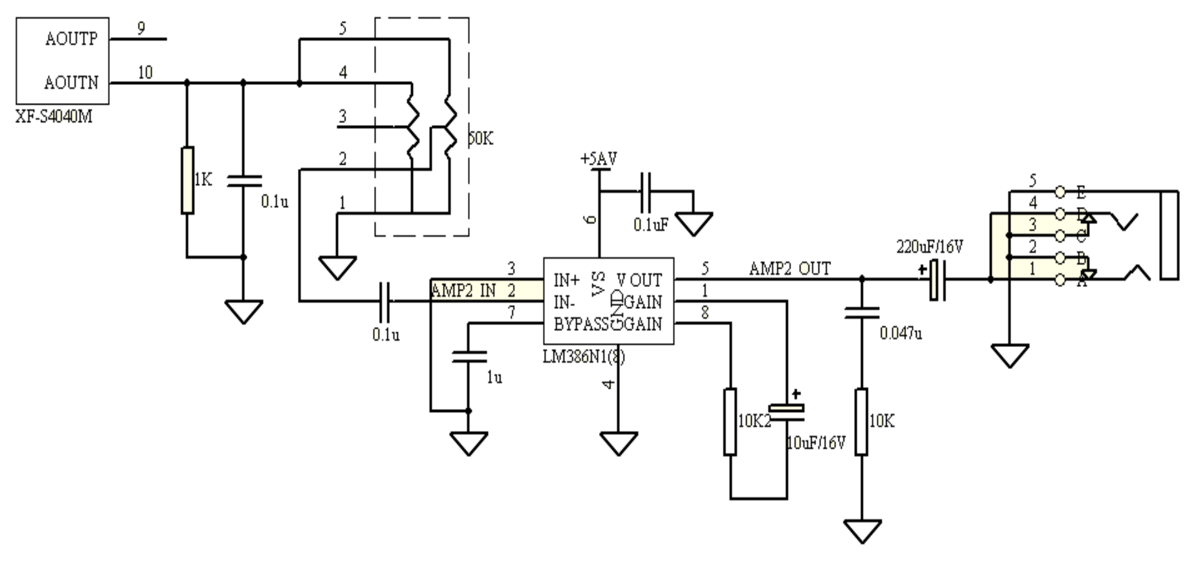


图表 16 XF-S4040 芯片与 MCU 的 I2C 接口连接示意图

9.4 音频功放电路



图表 17 XF-S4040 芯片差分音频输出与音频功放 TPA2005D1 的接口示意图



图表 18 XF-S4040 芯片与 LM386 的接口示意图

注： AOUTP 也可以直接接与 LM386 连接。接法和上述示意图一致

10. 极限值

图表 19 电气极限参数

参　　数	符　　号	最小值	最大值
电源	VDD(5V)	-0.5V	5.5V
电源	VDD(IO3V3) VDD(DAC3V3) VDD(DACREF)	-0.5V	3.6V
电源	VDD(CORE1V8)	-0.5V	1.95V
引脚输入	V_{IN}	-0.5	3.6V
工作温度（XF-S4040）	T_A	-25℃	+85℃
工作温度（XF-S4040M）	T_A	0℃	+70℃
储存温度	T_{STO}	-40℃	+125℃
静电防护	V_{ESD}	-2000V*1	+2000V*1
		-200V*2	+200V*2

*1 人体静电模型，等效放电模型为 1.5kohm 电阻对 100pF 电容。

*2 机器静电模型，等效放电模型为 0.75uH 电感与 10ohm 电阻串联对 200pF 电容。

图表 20 静态特性

参　　数	符 号	参　考　值			单位
		最小	典型	最大	
工作电压	VDD(5V)	3.7	5.0	5.5	V

工作电流（单 5V 供电）	I_{OP5V}	--	40	45	mA
工作电压	VDD(IO3V3) VDD(DAC3V3) VDD(DACREF)	3.0	3.3	3.6	V
工作电流 （1.8V 和 3.3V 分别供电）	I_{OP3V3}	--	5	8	mA
工作电压	VDD(CORE1V8)	1.7	1.8	1.95	V
工作电流 （1.8V 和 3.3V 分别供电）	I_{OP1V8}	--	35	40	mA
内部 3.3VLDO 输出*[1]	I_{LDO3V3}	--	25	30	mA
内部 1.8VLDO 输出*[1]	I_{LDO1V8}	--	50	60	mA
输入高电平	V_{IH}	2	--	3.6	V
输入低电平	V_{IL}	-0.5	--	0.8	V
输出电流	I_{OH}	--	-4	--	mA
吸收电流	I_{OL}	--	4	--	mA

*1　严禁使用 XF-S4040-CN 中文语音合成芯片内部的电压调节器给其他部件供电。

11. XF-S4040 语音合成芯片通讯协议

11.1 上位机发送的控制命令

上位机发送给 XF-S4040 的所有命令和数据都需要用“帧”的方式进行封装后传输。帧结构由帧头标志、数据区长度和数据区三部分组成。

图表 21 命令帧封装格式

帧头	数据区长度	数据区
0xFD(1Byte)	0xXX，0xXX（2Byte）	Data

图表 22 命令帧说明

名称	长度	说明
帧头	1 Byte	定义为十六进制“0xFD”
数据区长度	2 Bytes	用两个字节表示，高字节在前，低字节在后
数据区	小于 4096Bytes	命令字和命令参数，长度和“数据区长度”一致

数据区是由命令字和命令参数组成的，上位机使用命令字来实现语音合成芯片的各种功能。

图表 23 数据区命令字和命令参数

名称	发送的数据	说明
命令字	0x01	语音合成命令
	0x02	停止合成命令，没有参数

	0x03	暂停合成命令，没有参数
	0x04	恢复合成命令，没有参数
	0x21	TTS 系统状态查询命令
	0x88	系统进入 Power Down 模式，Reset 之后恢复
命令参数	不同命令字有不同参数列表，详见各命令字说明	

1. 语音合成命令

图表 24 语音合成命令

<table>
<tr><th>名称</th><th>发送的数据</th><th colspan="3">说明</th></tr>
<tr><td>命令字</td><td>0x01</td><td colspan="3">带文本编码设置的文本播放命令</td></tr>
<tr><td rowspan="6">参数列表</td><td rowspan="5">0xXX</td><td rowspan="5">1Byte 表示文本的编码格式，取值为 0～3</td><td>参数取值</td><td>文本编码格式</td></tr>
<tr><td>0x00</td><td>GB2312</td></tr>
<tr><td>0x01</td><td>GBK</td></tr>
<tr><td>0x02</td><td>BIG5</td></tr>
<tr><td>0x03</td><td>UNICODE</td></tr>
<tr><td>Data</td><td colspan="3">待合成文本的二进制内容</td></tr>
<tr><td>命令帧格式结构</td><td colspan="4">
<table>
<tr><td>帧头</td><td colspan="2">数据区长度</td><td colspan="3">数据区</td></tr>
<tr><td rowspan="2">0xFD</td><td>高字节</td><td>低字节</td><td>命令字</td><td>文本编码格式</td><td>待合成文本</td></tr>
<tr><td>0xHH</td><td>0xLL</td><td>0x01</td><td>0x00~0x03</td><td>…. ….</td></tr>
</table>
</td></tr>
<tr><td>示例</td><td colspan="4">
语音合成命令帧示例，发送文本编码格式为“GB2312”的文本“科大讯飞”
<table>
<tr><td>0xFD</td><td>0x00</td><td>0x0A</td><td>0x01</td><td>0x00</td><td>0xBF</td><td>0xC6</td><td>0xB4</td><td>0xF3</td></tr>
<tr><td>0xD1</td><td>0xB6</td><td>0xB7</td><td>0xC9</td></tr>
</table>
语音合成命令帧示例，发送文本编码格式为“GBK”的文本“科大讯飞”
<table>
<tr><td>0xFD</td><td>0x00</td><td>0x0A</td><td>0x01</td><td>0x01</td><td>0xBF</td><td>0xC6</td><td>0xB4</td><td>0xF3</td></tr>
<tr><td>0xD3</td><td>0x8D</td><td>0xEF</td><td>0x77</td></tr>
</table>
语音合成命令帧示例，发送文本编码格式为“BIG5”的文本“科大讯飞”
<table>
<tr><td>0xFD</td><td>0x00</td><td>0x0A</td><td>0x01</td><td>0x02</td><td>0xAC</td><td>0xEC</td><td>0xA4</td><td>0x6A</td></tr>
<tr><td>0xB0</td><td>0x54</td><td>0xAD</td><td>0xB8</td></tr>
</table>
语音合成命令帧示例，发送文本编码格式为“UNICODE”的文本“科大讯飞”
<table>
<tr><td>0xFD</td><td>0x00</td><td>0x0A</td><td>0x01</td><td>0x03</td><td>0xD1</td><td>0x79</td><td>0x27</td><td>0x59</td></tr>
<tr><td>0xAF</td><td>0x8B</td><td>0xDE</td><td>0x98</td></tr>
</table>
</td></tr>
<tr><td>特别说明</td><td colspan="4">当 XF-S4040 芯片正在合成文本的时候，如果又接收到一帧有效的数据，芯片会立即停止当前正在合成的文本，转而合成所接收到的最新的文本</td></tr>
</table>

2．停止合成命令

图表 25 停止合成命令

<table>
<tr><th>名称</th><th>发送的数据</th><th colspan="4">说明</th></tr>
<tr><td>命令字</td><td>0x02</td><td colspan="4">停止当前合成</td></tr>
<tr><td>参数列表</td><td>无</td><td colspan="4"></td></tr>
<tr><td rowspan="3">命令帧格式结构</td><td rowspan="2">帧头</td><td colspan="2">数据区长度</td><td colspan="2">数据区</td></tr>
<tr><td>高字节</td><td>低字节</td><td colspan="2">命令字</td></tr>
<tr><td>0xFD</td><td>0xHH</td><td>0xLL</td><td colspan="2">0x02</td></tr>
<tr><td>示例</td><td colspan="5">停止合成
0xFD 0x00 0x01 0x02</td></tr>
</table>

3．暂停合成命令

图表 26 暂停合成命令

<table>
<tr><th>名称</th><th>发送的数据</th><th colspan="4">说明</th></tr>
<tr><td>命令字</td><td>0x03</td><td colspan="4">暂停当前合成</td></tr>
<tr><td>参数列表</td><td>无</td><td colspan="4"></td></tr>
<tr><td rowspan="3">命令帧格式结构</td><td rowspan="2">帧头</td><td colspan="2">数据区长度</td><td colspan="2">数据区</td></tr>
<tr><td>高字节</td><td>低字节</td><td colspan="2">命令字</td></tr>
<tr><td>0xFD</td><td>0xHH</td><td>0xLL</td><td colspan="2">0x03</td></tr>
<tr><td>示例</td><td colspan="5">停止合成
0xFD 0x00 0x01 0x03</td></tr>
</table>

4．恢复合成命令

图表 27 恢复合成命令

名称	发送的数据	说明
命令字	0x04	恢复暂停的合成
参数列表	无	
命令帧格式结构	帧头 | 数据区长度（高字节 | 低字节） | 数据区（命令字） 0xFD | 0xHH | 0xLL | 0x04	
示例	停止合成 0xFD 0x00 0x01 0x04	

5．系统状态查询命令

图表 28 系统状态查询命令

名称	发送的数据	说明
命令字	0x21	通过该命令来判断 TTS 芯片是否正常工作，以及获取相应参数，返回 0x4E 表明系统仍在合成中，返回 0x4F 表明系统处于空闲状态
参数列表	无	
命令帧格式结构	帧头 | 数据区长度（高字节 | 低字节） | 数据区（命令字） 0xFD | 0xHH | 0xLL | 0x21	
示例	停止合成 0xFD 0x00 0x01 0x21	

6．进入 Power Down 状态命令

图表 29 进入休眠状态命令

<table>
<tr><th>名称</th><th>发送的数据</th><th>说明</th></tr>
<tr><td>命令字</td><td>0x88</td><td>进入 POWER DOWN 状态命令，RST 后恢复</td></tr>
<tr><td>参数列表</td><td>无</td><td></td></tr>
<tr><td>命令帧格式结构</td><td colspan="2"><table>
<tr><td>帧头</td><td colspan="2">数据区长度</td><td>数据区</td></tr>
<tr><td rowspan="2">0xFD</td><td>高字节</td><td>低字节</td><td>命令字</td></tr>
<tr><td>0xHH</td><td>0xLL</td><td>0x88</td></tr>
</table></td></tr>
<tr><td>示例</td><td colspan="2">停止合成
<table><tr><td>0xFD</td><td>0x00</td><td>0x01</td><td>0x88</td></tr></table></td></tr>
</table>

11.2 芯片反馈信息

XF-S4040 在初始化成功时会回传一个字节的“初始化成功”信息，初始化不成功上位机收不到此信息。在收到命令帧头“0xFD”后芯片会判断此命令帧正确与否，如果命令帧正确则返回“收到正确命令帧”回传，如果命令帧错误则返回“收到错误命令帧”回传。在芯片收到 TTS’系统状态查询命令时，如果 TTS 系统处于合成工作状态则返回“TTS 系统忙碌”信息，如果 TTS 系统处于空闲状态则返回“TTS 系统空闲”信息。在一帧数据合成完毕后，TTS 系统会自动返回一次“TTS 系统空闲”信息。

图表 30 芯片反馈信息

名称	发送的数据	触发条件
初始化成功回传	0x4A	系统初始化成功
收到正确的命令帧回传	0x41	收到正确的命令帧
收到错误命令帧回传	0x45	收到错误的命令帧
TTS 系统空闲	0x4F	1． 当一帧数据合成完以后，TTS 系统空闲状态回传 0x4F 2． 收到“状态查询命令”，TTS 系统处于空闲状态回传 0x4F
TTS 系统忙碌	0x4E	1． 收到“状态查询命令”，TTS 系统非空闲状态回传 0x4E

11.3 芯片工作状态引脚输出

图表 31 工作状态引脚描述

引脚序号	引脚名称	类型	描述
6	$\overline{Rdy}$/Bsy	输出	低电平：Ready 状态； 高电平：Busy 状态；

XF-S4040 芯片提供了一个硬件引脚指示芯片的工作状态，用户可以通过查询此状态输出引脚的电平，来判断芯片工作的状态。当此引脚处于低电平时，表明芯片处于空闲状态，没有合成文本；当引脚处于高电平状态时，表明芯片处于合成文本的工作状态。

11.4 上位机对 XF-S4040 芯片调用方式

11.4.1 简单调用

简单调用针对初级用户或应用比较简单的情况。用户不用关心芯片的工作状态，只需要发送文本，芯片会将接收的文本合成为语音输出。

在简单调用情况下，上位机只要与芯片之间建立起 UART 或者 I^2C 两种通信方式中的任意一种，即可发送合成命令来实现文本的合成，上位机不需要理睬芯片的反馈信息和状态输出，XF-S4040 立即输出合成的语音。

提示：如前一帧文本还没有合成完，再发送文本到芯片就会打断前次合成，而执行新的合成。

11.4.2 标准调用方式

对于一般情况，上位机需要了解芯片的工作状态，以更精确的控制芯片的动作：比如需要确保上次文本被完整合成之后，再合成下一段文本。为了实现这个目的，有三种实现方法：

1. 通过查询 $\overline{Rdy}$/Bsy 引脚来了解芯片的工作状态，如此引脚处于高电平状态，说明芯片正在处于文本合成状态，需要继续等待，直到此引脚变成低电平状态后，系统处于空闲状态，再给芯片发送下一段文本进行合成；
2. 向芯片发送系统状态查询命令，通过返回值即可知道系统是否空闲（详见图表 27）；
3. 接收来自芯片的反馈信息，当文本合成完毕后，芯片会反馈“0x4F”字符，表示系统已经空闲（对于 UART 通讯方式上位机可以直接接收此信息，对于 SPI 和 I^2C 通讯方式，由于芯片是 Slave 身份，需要上位机主动读取此信息）；

应用举例如下：假设需要合成的文本为 5000 字节，超过了芯片一个命令帧所能容纳的最大文本长度（4096 字节），这时分两次给芯片发送文本信息。程序过程如下：

1、 上位机先给芯片发送一个文本合成命令帧，携带 4096 个字节的文本；
2、 上位机等待芯片返回播放完毕回传信息，直到收到芯片反馈“0x4F”系统空闲信息；
3、 上位机给芯片再次发送一个文本合成命令帧，发送出剩下的 4 个字节文本信息。